

Study	Patient selection				Test integrity								Data Flow			
	Consecutive or random sample	case-control avoided/avoid artificially inflated base rate	Avoided inappropriate exclusions	Overall risk of bias	PHQ-9 blinding to reference test	if index test translated, appropriate translation?	if translated, psychometric properties reported	Index test: overall risk of bias	Reference test correctly classifies target condition	if reference test translated, appropriate translation	if translated, psychometric properties reported	Reference test: overall risk of bias	Flow/timing: interval of 2 weeks or less	Flow/timing: all participants receive same reference test	Flow/timing: > 80% participants included in analysis?	Flow/ timing: overall risk of bias
<i>Arroll et al 2010</i>	✓	✓	✓	Low	✓	✓	n/a	n/a	Low	n/a	✓	n/a	n/a	Low	✓	✓
<i>Ayalon et al 2010</i>	✓	✓	✓	Unclear		?	✓	?	Unclear	✓	?	✓	✓	Unclear		?
<i>Azah et al 2005</i>	?	?	✓	Unclear	✓	?	✓	x	high	✓	?	?	?	high	?	✓
<i>Cannon et al 2007</i>	✓	x	?	High	✓	✓	n/a	n/a	low	n/a	✓	n/a	n/a	Low	✓	?
<i>Chen et al 2013</i>	✓	✓	✓	Low	✓	✓	✓	?	Unclear	✓	✓	n/a	n/a	low	✓	✓
<i>Chen et al 2010</i>	✓	✓	✓	Low	✓	✓	?	?	Unclear	?	✓	?	?	Unclear	✓	✓
<i>Osório et al 2009</i>	✓	✓	✓	Low	X	?	✓	x	High	✓	?	✓	x	high	✓	✓
<i>Gelaye et al 2013</i>	x	x	?	high	✓	✓	✓	?	Unclear	✓	✓	?	?	Unclear	✓	✓
<i>Gilbody et al 2007</i>	✓	✓	✓	Low	✓	?	n/a	n/a	Unclear	n/a	?	n/a	n/a	Unclear	✓	✓
<i>Henkel et al 2004</i>	x	✓	✓	Low	✓	?	n/a	n/a	Unclear	n/a	?	n/a	n/a	Unclear	✓	✓
<i>Kroenke et al 2001</i>	✓	✓	✓	Low	✓	✓	n/a	n/a	low	n/a	✓	n/a	n/a	Low	X	✓
<i>Lamers et al 2008</i>	x	x	x	high	✓	✓	?	?	Unclear	?	?	?	?	Unclear	X	?
<i>Liu et al 2011</i>	✓	✓	✓	Low	✓	✓	✓	?	Unclear	✓	✓	✓	?	Unclear	✓	✓
<i>Lotraku et al 2008</i>	x	✓	?	high	✓	✓	✓	?	Unclear	✓	✓	✓	✓	Low	✓	?
<i>Lowe et al 2005</i>	✓	✓	✓	Unclear	✓	✓	n/a	n/a	Low	n/a	✓	n/a	n/a	Low	✓	?

BJPO/2015/001685 **Data supplement**
doi: 10.1192/bjpo.bp.115.001685

<i>Lowe et al 2004</i>	✓	✓	✓	Low	✓	✓	n/a	n/a	Low	n/a	✓	n/a	n/a	Low	X	✓
<i>Patel et al 2007</i>	✓	✓	?	Unclear	✓	✓	✓	?	Unclear	✓	✓	✓	?	unclear	✓	?
<i>Phelan et al 2010</i>	✓	✓	✓	Low	✓	✓	n/a	n/a	Low	n/a	✓	n/a	n/a	Low	✓	✓
<i>Richardson et al</i>	✓	✓	✓	Low	X	✓	n/a	n/a	Low	n/a	✓	n/a	n/a	Low	✓	?
<i>Sherina et al 2012</i>	✓	✓	✓	Low		✓	✓	?	Unclear	✓	✓	✓	?	unclear		✓
<i>Spitzer et al 1999</i>	x	✓	✓	high	✓	✓	n/a	n/a	Low	n/a	✓	✓	n/a	Low	✓	✓
<i>Sung et al 2013</i>	✓	✓	✓	Low		?	?	?	Unclear	?	?	?	?	unclear		?
<i>Wittkamp et al 2009</i>	?	x	x	High	✓	✓	n/a	n/a	Low	n/a	✓	n/a	n/a	Low	X	?
<i>Yeung et al 2008</i>	?	✓	✓	Low	X	✓	✓	?	Unclear	✓	✓	?	?	Unclear	X	?
<i>Zuithoff et al 2010</i>	✓	✓	✓	Low	✓	✓	✓	?	Unclear	✓	✓	✓	✓	Low	✓	?

Footnote: ✓ = satisfactory (criterion met); X, not satisfactory (criterion not met); ?, unclear (unable to establish if criterion met); n/a, not applicable.